Evaluation of a Belief-based Decision Making in a Real-time Platform for Cognitive Radio Networks

A. Raschellà, J. Pérez-Romero, O. Sallent, A. Umbert

Dept. of Signal Theory and Communications
Universitat Politècnica de Catalunya (UPC)
Barcelona, Spain
e-mail: [alessandror, jorperez, sallent, annau]@tsc.upc.edu

Abstract—**This paper presents a real-time testbed based on reconfigurable devices able to transmit and receive data at different operating frequencies, which are dynamically configured. The main objective of this paper is to provide an advanced and realistic framework where the performance of algorithms and policies for spectrum selection in Cognitive Radio (CR) networks can be fully evaluated in a real-time environment. In particular, this paper focuses on a CR belief-based decision making framework particularized for different observation strategies that decide when measurements are carried out to predict the radio environment dynamism. Results show that an accurate selection of an observation strategy is a key element for the trade-off between achieved performance and measurement requirements.**

Keywords-cognitive radio; testbed; spectrum selection; decision making; belief vector.

I. INTRODUCTION

Cognitive Radio (CR) has been identified as a promising solution to solve the conflicts between spectrum demand growth and spectrum underutilization [1][2]. In fact, CR paradigm enables flexible, efficient and reliable spectrum use by adapting the radio operating characteristics to the real-time conditions of the environment. The operation of a CR system is typically assumed to follow the cognitive cycle that gathers observations from the outside world through different types of sensors, makes decisions between the different choices, and then executes the decisions made through actions.

Hence, one important function within the CR cycle is the observation stage, which typically is in charge of making measurements at several nodes of a CR network. Then, these measurements need to be reported to the node in charge of the decision making through signaling procedures supported by cognitive control channels [3]. Consequently, the observation stage can be very costly in terms of signaling overhead, battery consumption, etc. Therefore, decision making strategies able to efficiently operate with the minimum amount of measurements become of high interest for enhancing CR operation.

In the literature different decision making tools have been proposed with the aim to operate with reduced measurement needs. For instance, several research works rely on Partially Observable Markov Decision Processes (POMDPs) [4] as a decision making tool that combines partial observations of the radio environment at specific periods of time with a statistical characterization of the system dynamics. For instance, [5] and [6] propose opportunistic spectrum access approaches to channels that can be either busy or idle, assuming a single unlicensed user. In [7] the problem was studied for a multi-user scenario through a collaborative approach. In [8][9] the use of POMDP for spectrum selection in CR networks is proposed. Some other works rely on the restless multi-armed bandit problem to find the optimal policy of sensing channels so as to maximize the expected throughput [10]. This paper extends the mentioned works by exploiting the use of the so-called *belief vector* to predict the environment dynamics. In details, the belief vector assesses the probability that the radio environment is under specific conditions (e.g. interference levels) at a certain time based on past measurements. As long as the belief vector predicts with sufficient accuracy the existing conditions at the decision making time, smart and proper decisions can be made with minimum requirements in terms of observations. In that context, this paper proposes a general formulation of the belief-based decision making particularized for different observation strategies.

The frameworks presented in the abovementioned works rely mainly on simulation or analytical studies. While this is a usual approach in the research community, it involves some simplifications of the real environment so it can be useful for obtaining preliminary results. Nevertheless, to conduct meaningful and appropriate studies, and to accurately assess the performance of innovative solutions, the evaluation over more realistic platforms is becoming essential as a step forward towards the implementation in a real system.

Therefore, the belief-based decision making proposed in this paper has been designed, developed and evaluated by means of a real-time testbed consisting of a hardware platform and a software component. The developed platform is a powerful tool where the benefits of the proposed framework can be demonstrated in a real time environment. A preliminary version of the testbed with basic functionalities can be found in [11].

The rest of the paper is organized as follows. Section II illustrates the considered system model and formulates the general belief-based decision making approach. This general approach is particularized for different observation strategies presented in Section III. Section IV illustrates the testbed implementation while, Section V presents the model implemented in the testbed to evaluate the proposed approach

as well as some performance results. Finally, Section VI points out concluding remarks and future works.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The system model considered in this paper assumes a set of $j=1,..., L$ radio links, each one intended to support data transmission between either a pair of terminals or between a terminal and an infrastructure node. The j–th radio link will be supporting a certain data service characterized by a required bit rate $R_{req,j}$ and will be generating data transmission sessions of a certain duration D_j.

The potential spectrum to be assigned to the different radio links is organized in a set of $i=1,..., M$ spectrum blocks (SBs). Each one is characterized by a central frequency and bandwidth. In general, the SBs can belong to different spectrum bands subject to different interference conditions and to different regulatory regimes (e.g., unlicensed bands, shared bands, licensed bands, etc.). Interference in each SB is assumed to come from other transmitters that have the same rights of use than the considered radio links (e.g., in case of shared bands, such as TVWS, it is assumed that the interference comes from other secondary devices and there are no primary users in the area where the links operate).

The available bit rate for the j-th link in the i-th SB $R_{j,i}$ will depend on both the propagation conditions between the j-th link transmitter and receiver as well as on the interference in the i-th SB experienced at the receiver. Then, the problem considered here consists in performing an efficient allocation of the SBs to the radio links by properly matching the bit rate requirements with the achievable bit rate in each SB. For that purpose, it is assumed that the different radio links are controlled by a centralized decision making entity residing at the infrastructure side in charge of deciding the spectrum to be used by each radio link. Some illustrative use cases where this system model can be applicable are: (i) a Digital Home scenario in which different devices need to communicate, (ii) a set of cognitive small cells deployed in a cellular network that make use of additional spectrum to increase the network capacity, and (iii) an opportunistic Device-To-Device (D2D) radio link created to extend the coverage of certain cellular terminals that are outside the coverage area of the cellular infrastructure.

The spectrum selection decision-making will take a so-called *action*, corresponding to the allocation of a SB to a radio link, anytime that a data transmission session is initiated on this radio link. The action made for the j–th link at time t is denoted as $a_j(t) \in \{1,..., M\}$ and corresponds to the selected SB among those currently available.

The considered interference model denotes as $I_{j,i}(t)=I_{max,j,i} \cdot \sigma_i(t)$ the interference spectral density measured by the receiver of the j-th link in the i-th SB at a given time due to other external transmitters (i.e. not belonging to the L radio links). In order to capture that interfering sources may exhibit time-varying characteristics, $\sigma_i(t)$ is a SB-specific term between 0 and 1 (i.e. $\sigma_i(t)=0$ when no interference exists and $\sigma_i(t)=1$ when the interference reaches its maximum value $I_{max,j,i}$). For modelling purposes, it is considered that the set of possible values of $\sigma_i(t)$ is translated into a discrete set of

interference states $S^{(i)}(t) \in \{0,1,..., K\}$ where state $S^{(i)}(t)=k$ corresponds to $\sigma_{k-1} < \sigma_i(t) \le \sigma_k$ for $k>0$ and to $\sigma_i(t)=\sigma_0=0$ for $k=0$. Note also that $\sigma_K=1$.

The interference evolution for the i-th block is modelled as an ergodic discrete-time Markov process with the state transition probability from being in state k at time t and moving to state k' in the next time step $t+1$ given by:

$$p_{k,k'}^{(i)} = \Pr\left[S^{(i)}(t+1) = k' \,\middle|\, S^{(i)}(t) = k \right] \qquad (1)$$

It is assumed that the state of the i-th SB, $S^{(i)}(t)$, evolves independently from the other SBs, and that the state evolution is independent from the assignments made by the spectrum selection algorithm. Note that, without loss of generality, the time axis is assumed to be discretized in time steps. Then, the state transition probability matrix for the i-th SB is defined as $\mathbf{P^{(i)}}$ where the element in the k-th row and k'-th column corresponds to $p_{k,k'}^{(i)}$.

Moreover, let us define the steady state probability vector $\boldsymbol{\pi^{(i)}}$ as a column vector, whose k-th component $\pi_k^{(i)}$ is the probability that the i-th SB is in the k-th interference state.

Each radio link with a data session in course (referred to as an active link) will obtain a reward that measures the obtained performance depending on the interference state of the allocated SB at each time. $r_{j,k}^{(i)}$ denotes the reward that the j-th link gets when using its allocated SB i and the interference state is $S^{(i)}(t)=k$. The reward is a metric between 0 and 1 capturing how suitable the i-th SB is for the j-th radio link, depending on the bit rate that can be achieved in this SB with respect to the bit rate required by the application $R_{req,j}$. The reward vector of the j-th link in the different interference states of the i-th SB is defined as the column vector $\mathbf{r_j^{(i)}}$ whose k-th component is $r_{j,k}^{(i)}$. It is worth mentioning that many possible definitions of the reward metric as a function of the bit rate may exist (e.g. sigmoid functions, linear functions, etc.).

The average reward experienced on the j-th link and i-th SB along a session starting to transmit data at time $t+1$ and ending after a certain duration D_j time steps, is given by:

$$r_{SESSION,j}^{(i)} = \frac{1}{D_j} \sum_{n=1}^{D_j} r_{j,S^{(i)}(t+n)}^{(i)} \qquad (2)$$

With all the above foundations, the spectrum selection policy executed at time t for the j-th radio link will target the maximization of the expected reward that the session will experience along its duration:

$$a_j(t) = \arg \max_{\substack{i \in \{1,...,M\} \\ i \text{ available}}} E\left[\frac{1}{D_j} \sum_{n=1}^{D_j} r_{j,S^{(i)}(t+n)}^{(i)} \right] \qquad (3)$$

where the selection is made among the subset of available SBs, i.e. those that are not allocated to any other radio link at the *decision* making time t.

The *analysis* of the future evolution of the reward in each of the SBs until the session end will exploit measurements (observations) of the interference state of the different SBs

carried out at specific time instants in the past, together with the statistical characterisation of the interference dynamics in each SB. In particular, denoting as $o^{(i)}(t-m^{(i)})$ the *observation* (measurement) that provides the value of the interference state of the i-th SB that was measured at time step $t-m^{(i)}$, i.e. $o^{(i)}(t-m^{(i)})=S^{(i)}(t-m^{(i)})$, the criterion of (3) can be reformulated in order to exploit knowledge from the observations in the past as:

$$a_j(t) = \arg \max_{\substack{i\in\{1,\dots,M\} \\ i\ \text{available}}} \Phi_j^{(i)}(t) \tag{4}$$

where $\Phi_j^{(i)}(t)$ is the SB-dependent decision function to be maximized, given by:

$$\Phi_j^{(i)}(t) = \frac{1}{\overline{D_j}}\sum_{n=1}^{\overline{D_j}} E\left[r_{j,S^{(i)}(t+n)}^{(i)} \Big| o^{(i)}\left(t-m^{(i)}\right) = S^{(i)}\left(t-m^{(i)}\right)\right] \tag{5}$$

Notice that, since the session duration D_j will usually be random and unknown at the decision making time t, it has been characterised in (5) statistically in terms of its average value $\overline{D_j}$. In (5), the estimation of the expected reward achieved in the i-th SB at future time instants $t+n$ based on the past observation at $t-m^{(i)}$ will rely on the statistical characterization of the interference dynamics given by the so-called *belief vector*. It is defined as vector $\mathbf{b}^{(i)}(t)$ where component $b_k^{(i)}(t)$ is the conditional probability that the i-th block will be in state $S^{(i)}(t)=k$ at time t given the last observation of the actual interference state that was taken at time step $t-m^{(i)}$, that is:

$$b_k^{(i)}(t) = \Pr\left[S^{(i)}(t)=k \Big| o^{(i)}\left(t-m^{(i)}\right) = S^{(i)}\left(t-m^{(i)}\right)\right] \tag{6}$$

Then, the expected reward obtained in the i-th SB at time $t+n$ can be expressed in terms of the belief vector as:

$$E\left[r_{j,S^{(i)}(t+n)}^{(i)} \Big| o^{(i)}\left(t-m^{(i)}\right) = S^{(i)}\left(t-m^{(i)}\right)\right] = \mathbf{b}^{(i)\mathbf{T}}(t+n)\,\mathbf{r_j^{(i)}} \tag{7}$$

where superscript T denotes transpose operation. By making use of (7) the estimation of the average reward achieved in the i-th SB along the session duration of the j-th radio link given by decision function (5) can be expressed as:

$$\Phi_j^{(i)}(t) = \frac{1}{\overline{D_j}}\sum_{n=1}^{\overline{D_j}} \mathbf{b}^{(i)\mathbf{T}}(t+n)\,\mathbf{r_j^{(i)}} \tag{8}$$

The computation of the belief vector of the i-th SB at a certain time instant t is done recursively starting from the last observation of the actual interference state that was taken at time step $t-m^{(i)}$. Specifically, considering that $o^{(i)}(t-m^{(i)})=S^{(i)}(t-m^{(i)})$ the components of the belief vector at time $t-m^{(i)}$ are given by:

$$b_k^{(i)}\left(t-m^{(i)}\right) = \begin{cases} 1 & \text{if } k=S^{(i)}\left(t-m^{(i)}\right) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

This is expressed in vector notation as:

$$\mathbf{b}^{(i)}\left(t-m^{(i)}\right) = \mathbf{x}\left(S^{(i)}\left(t-m^{(i)}\right)\right) \tag{10}$$

where $\mathbf{x}(k)$ is defined as a column vector of $K+1$ components numbered from 0 to K that has all of them equal to 0 except the k-th component that is equal to 1.

Then, the belief vector at a time instant $t > t-m^{(i)}$ can be obtained from the belief vector at the previous time step t - 1 making use of the state transition probability matrix as:

$$\mathbf{b}^{(i)\mathbf{T}}(t) = \mathbf{b}^{(i)\mathbf{T}}(t-1)\cdot\mathbf{P^{(i)}} \tag{11}$$

By recursively applying (11) for the last $m^{(i)}$ time steps and by making use of (10) the belief vector at time t as a function of the last observation is given by:

$$\mathbf{b}^{(i)\mathbf{T}}(t) = \mathbf{b}^{(i)\mathbf{T}}\left(t-m^{(i)}\right)\left[\mathbf{P^{(i)}}\right]^{m^{(i)}} = \mathbf{x}^{\mathbf{T}}\left(S^{(i)}\left(t-m^{(i)}\right)\right)\left[\mathbf{P^{(i)}}\right]^{m^{(i)}} \tag{12}$$

III. Observation Strategies

The observation strategy specifies the time instants when the actual interference state in each SB is measured. The observation strategy should make sure that the time $m^{(i)}$ elapsed between the last observation and the spectrum selection decision making time t is adequate enough to compute the belief vector $\mathbf{b}^{(i)\mathbf{T}}(t+n)$ and make accurate decisions based on (8). Then, this paper considers the following three observation strategies:

a) *Instantaneous Measurements (IM) strategy*: It consists in performing instantaneous measurements of the interference states in all the SBs at the time t when a new session has to be established, i.e. at the time when the spectrum selection decision-making is executed. In this case, the belief vector will always be computed with $m^{(i)}=0$ and therefore it will capture the exact interference state at time t. Then, the belief vector at t will be given by:

$$\mathbf{b}^{(i)\mathbf{T}}(t) = \mathbf{x}^{\mathbf{T}}\left(S^{(i)}(t)\right) \tag{13}$$

Correspondingly, the decision function becomes:

$$\Phi_j^{(i)}(t) = \frac{1}{\overline{D_j}}\mathbf{x}^{\mathbf{T}}\left(S^{(i)}(t)\right)\left(\sum_{n=1}^{\overline{D_j}}\left[\mathbf{P^{(i)}}\right]^n\right)\mathbf{r_j^{(i)}} \tag{14}$$

b) *Periodic Measurements (PM) strategy*: It consists in performing periodic measurements of the i-th SB with observation period $T_{obs}^{(i)}$. In this way, the elapsed time $m^{(i)}$ between the last observation of the i-th SB and the decision making time t will always be upper bounded by $m^{(i)} \leq T_{obs}^{(i)}$. It will be assumed that only the SBs that are not allocated to any link will be measured, since they are the only SBs that can be considered in the decision making process. In turn, when a SB is released, it will also be measured in case that the time since the last observation exceeds $T_{obs}^{(i)}$.

c) *Steady-state (StS) strategy*: This is the simple case in which no actual observations are performed. In this case, $m^{(i)} \to \infty$ and it can be easily proved making use of the properties of ergodic discrete time Markov processes [12] that the values of the belief vector will be equal to the steady-state probabilities $\boldsymbol{\pi}^{(i)}$. Then the decision function becomes:

$$\Phi_j^{(i)}(t) = \lim_{m^{(i)} \to \infty} \frac{1}{\overline{D}_j} \sum_{n=1}^{\overline{D}_j} \mathbf{b}^{(i)\mathbf{T}}(t+n)\mathbf{r}_j^{(i)} = \boldsymbol{\pi}^{(i)\mathbf{T}}\mathbf{r}_j^{(i)} \quad (15)$$

IV. TESTBED IMPLEMENTATION

This section describes the architecture of the platform developed to test the proposed framework together with the details about the hardware and software components.

A. Hardware Component

The testbed platform is made of several nodes implemented through Universal Software Radio Peripheral (USRP) [13] integrated boards controlled by a Personal Computer (PC) running Linux operating system. USRP is a Software Defined Radio (SDR) hardware platform that provides radio front-end functionalities, Analog to Digital and Digital to Analog Conversion (ADC/DAC), a Radio Frequency (RF) front end, a Field Programmable Gate Array (FPGA) which executes some pre-processing of the input signal and a USB 2.0 interface to connect to a PC. On the motherboard, there are four slots, where up to 2 RX and 2 TX daughterboards can be plugged in.

The daughterboards hold the RF receiver and the RF transmitter. There are 4 high-speed 12-bit ADCs and 4 high-speed 14-bit DACs. All the ADCs and DACs are connected to the FPGA that performs high bandwidth math procedures such as filtering, interpolation and decimation. The DACs clock frequency is 128 Msample/s, while ADCs work at 64 Msample/s to digitize the received signal. A USB controller sends the digital signal samples to the PC in 16-bit I and 16-bit Q complex format (4 bytes per complex sample), resulting in a maximum rate of 8 Msample/s per daughterboard. Consequently, the FPGA has to perform filtering and digital down-conversion (decimation) to adapt the incoming data rate to the USB 2.0 and PC computing capabilities. Therefore, the maximum RF bandwidth that can be handled is 8 MHz (narrower bandwidths can be selected by adjusting the decimation rate).

There exist different kinds of daughterboards that allow a very high USRP reconfigurability and working at several frequency bands. In this testbed each USRP consist of two daughterboards *XCVR2450 Transceivers* working in the frequency ranges 2.4 - 2.5 GHz and 4.9 - 5.9 GHz.

B. Software Component

Each PC runs the GNU radio software [14], a free and open source toolkit that provides a library of signal processing blocks for building SDRs. In GNU radio, the programmer builds a radio by creating a graph where the vertices are signal processing blocks and the edges represent the data flow between them. All the signal processing blocks are written in C++ and Python is used to create a network or graphs and glue these blocks together. Simplified Wrapper and Interface Generator (SWIG) is an open source package used by GNU radio as glue such that the C++ classes can be used from Python. SWIG has the ability to convert the C++ classes into Python compatible ones. As a result, the whole GNU radio framework is capable of putting together and exploiting the benefits of both C++ and Python. GNU radio has been used to develop the formulation of the belief-based decision making

described in the previous section and to enable the data and control communication between USRP transceivers.

C. Testbed Architecture

The scenario where the proposed framework has been implemented is illustrated in Fig. 1.
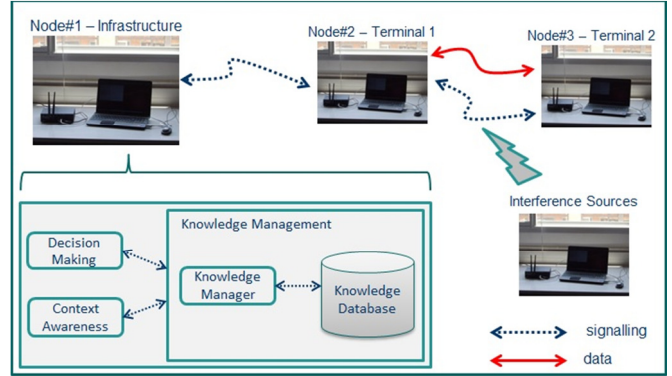

Figure 1. Testbed Architecture

Node#1 represents the infrastructure side in charge of the following main functionalities:

1) Decision-Making entity: it is in charge of selecting the most appropriate SB following the decision function (8).

2) Knowledge Management entity: it includes the Knowledge Database (KD) and the Knowledge Manager (KM). The KD stores the information about state transition matrices, steady-state probabilities, reward values, belief vectors and average session durations. The KM is in charge of computing the belief vector following (12) based on the information in the KD.

3) Context Awareness (CA) entity: it is in charge of performing measurements of the interference states of the different SBs. This is needed for acquiring the statistics of the state transition matrices and the steady-state probabilities stored in the KD and for dynamically tracking the interference variability in the different SBs during the system operation based on the observation strategies discussed in Section III. Measurements are performed by means of the energy detection functionality implemented in the USRPs. Energy detection for each SB is performed during a sensing time Δt_m. Then, based on the detected energy, the CA identifies the interference states of each SB. The energy threshold to decide if a SB is free of interference is set based on [15].

Node#2 and Node#3 in Fig. 1 represent the terminals that need to establish the link for data transmission. Node#2 is programmed to compute periodically the reward that measures the obtained performance depending on the experienced bit rate in the allocated SB. The reward follows the formulation defined in [8] ranging from 0 to 1, where 1 corresponds to the case when the obtained bit rate equals the required value R_{req}.

Moreover, three USRP-based interference sources controlled by PCs have been included in the proposed platform. These interference sources are transmitting in specific SBs following random patterns whose statistics can be controlled at the testbed configuration.

V. PERFORMANCE EVALUATION

A. Scenario parameters

This section describes the emulation assumptions that have been considered to evaluate the performance achieved by the proposed framework. A set of $M=3$ SBs are taken into account. The SB bandwidths are 200 kHz and the central frequencies are 5472, 5490, 5508 MHz. Two different interference states are considered for the SBs: $S^{(i)}=0$ when no interference exists and $S^{(i)}=1$ when the interference corresponds to its maximum value. The average durations of the interferences states for each SB are presented in Table I. $L=1$ link is considered to transfer the data flow between the terminals with bit rate requirement $R_{req}=$ 512 kbps. The testbed operates in time steps of 10 seconds. The link session duration is exponentially distributed with average $\overline{D}_j=$ 15 time steps and the time between the end of a session and the beginning of the next one is 1 time step.

TABLE I. DURATIONS OF THE INTERFERENCE STATES FOR THE SBS

State	SB1	SB2	SB3
$S^{(i)}=0$	480 steps	60 steps	480 steps
$S^{(i)}=1$	120 steps	480 steps	160 steps

B. State transition probability acquisition

A first study has been carried out in order to evaluate the performance of the measurement procedure implemented in the infrastructure node in order to acquire the values of the state transition probabilities to be stored in the KD. For each SB the estimation of these probabilities is done by sensing the interference state once per time step of 10s. This is used to dynamically track the time that the SB remains in each state, from which the state transition probabilities can be easily computed. Fig. 2 presents the time evolution of the estimated state transition probabilities $p_{0,0}^{(1)}$ and $p_{1,1}^{(1)}$ in SB1 for two different values of the sensing time duration $\Delta t_m =$2s and Δt_m =10s used by the CA. The theoretical values according to the average durations of the interference state for SB1 defined in Table I are also included.
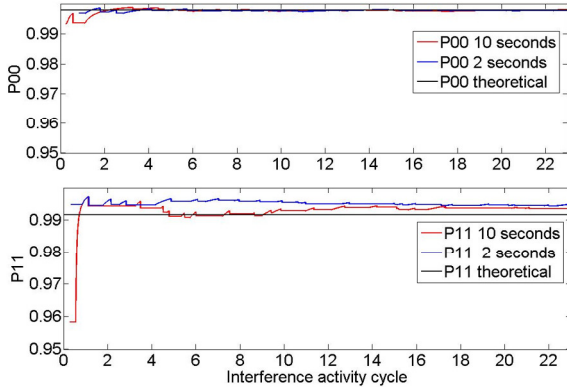

Figure 2. Evolution of the estimated state transition probabilities of SB1 based on measurements

Note that in order to estimate the state transition probabilities, the CA needs to have observed a certain number of full periods in state $S^{(1)}=1$ and a certain number of full periods in state $S^{(1)}=0$. For this reason, the time axis in Fig. 2 is expressed in terms of the number of *interference activity*

cycles, where the duration of one cycle corresponds to the average duration in state $S^{(1)}=0$ plus the average duration in state $S^{(1)}=1$, i.e. 600 time steps. The total duration of the study is 23 cycles that correspond to 38.3 h. From the figure it can be observed that after few cycles the estimated values of $p_{0,0}^{(1)}$ and $p_{1,1}^{(1)}$ obtained based on measurements converge to the expected ones with very low error (i.e. 0.01% in case of $p_{0,0}^{(1)}$ and 0.2% in case of $p_{1,1}^{(1)}$). Moreover, no significant differences are observed for the two considered values of Δt_m.

C. Performance results of the observation strategies

The assessment of the proposed framework has been carried out in terms of different Key Performance Indicators (KPIs). In particular, the *average reward* experienced by each data transmission session depending on the interference state of the allocated SB and averaged along the total emulation time is considered first. In addition, the *average throughput* defined as the achieved bit rate averaged along the total emulation time is also evaluated. Finally the testbed also measures the *observation rate* defined as the average number of observations per step that are performed to determine the interference states of the SBs. This KPI is only applicable to IM and PM policies, while StS strategy does not require observations of the system.

Table II presents a comparison between the different observation strategies discussed in Section III in terms of the abovementioned KPIs obtained after an emulation time $T_{em}=$ 3600 time steps and with Δt_m=2s. PM strategy is configured with $T_{obs}^{(i)}$=50 time steps. As an additional baseline reference, the random algorithm is also considered, in which the SB is selected randomly among the available ones without making any observation. In addition, Fig. 3 presents the time evolution of the different strategies in terms of average reward.

TABLE II. PERFORMANCE RESULTS ACHIEVED WITH THE DIFFERENT STRATEGIES

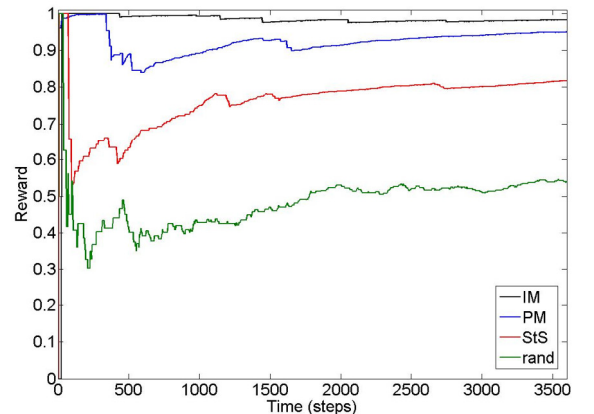| KPI	IM	PM	StS	random
Reward	0.98	0.95	0.81	0.54
Th. (kbps)	502	486	413	305
Obs. Rate (obs/step)	0.19	0.07	0.0	0.0

Figure 3. Evolution of the average reward

Firstly it can be observed from Fig. 3 and Table II how all the proposed solutions allow achieving a clear improvement in

terms of both reward and throughput with respect to the random selection of the SB. Moreover, it can be noticed that PM and IM strategies achieve performance improvements of around 20% with respect to StS. Hence, taking into account that PM policy allows a reduction of approximately 60% with respect to IM in terms of observation rate (see Table II), such a solution is a good trade-off between efficient performance and reduction of measurement needs.

Fig. 4 and Fig. 5 illustrate the performance in terms of reward and observation rate among the different strategies as a function of the observation period $T_{obs}^{(i)}$ used by PM (hence, $T_{obs}^{(i)}$ values do not affect the performance of the other strategies). From Fig. 4 it can be observed that the performance achieved by PM is a decreasing function of the observation period. In particular, low values such as $T_{obs}^{(i)}$=20 time steps lead to a large reward at the expense of an increase in the observation rate. In turn, when increasing the value of $T_{obs}^{(i)}$, the observation rate can be substantially reduced but with a degradation in terms of reward. Hence, to take advantage of the good trade-off between reward/observation rate achieved by the PM strategy, a value of $T_{obs}^{(i)}$ approximately between 50 and 100 time steps would be appropriate in the scenario considered in this paper.
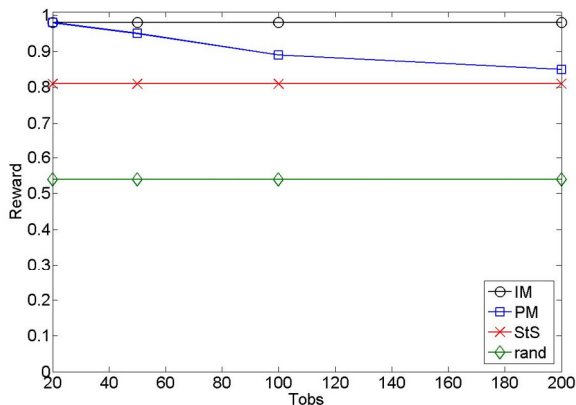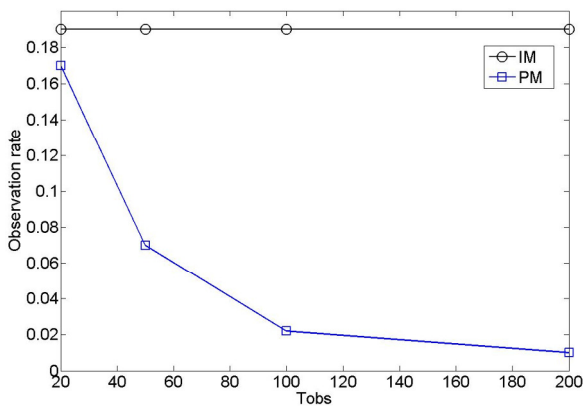

Figure 4. Reward results as a function of T$_{obs}$

Figure 5. Observation rate results as a function of T$_{obs}$

VI. CONCLUSIONS

This paper has presented a real-time testbed based on USRP hardware platform and GNU radio software. It represents a realistic framework to develop and validate algorithms and policies for spectrum selection in CR networks. The testbed has been used to evaluate a proposed decision making solution based on the belief vector concept. It allows predicting the environment dynamics at the decision making time and in later instants based on past measurements. Results have illustrated that an appropriate combination between periodical measurements of the radio environment and statistical characterization of the interference variations is a good trade-off between high performance and reduction of measurement needs in the scenario considered in this paper.

Future work will extend the study proposed in this paper considering the impact of the dynamicity of the radio environment in terms of the interference states durations and of the traffic generation patterns, in order to better assess the convenience of one or another observation strategy in different scenarios.

REFERENCES

[1] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. dissertation, KTH Royal Institute of Technology, 2000.

[2] J. Vartiainen, M. Hoyhtya, J. Lehtomaki, and T. Braysy, "Priority channel selection based on detection history database" CROWNCOM 2010, Cannes, France, June 2010.

[3] V. Stavroulaki, et al. "Cognitive Control Channels: From Concept to Identification of Implementation Options", IEEE Communications Magazine, Vol. 50, No. 3, pp. 96-108, July 2012.

[4] K.P. Murphy, "A Survey of POMDP Solution Techniques", available at http://http.cs.berkeley.edu/~murphyk/Papers/pomdp.ps.gz, 2000.

[5] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework", IEEE Journal on Selected Areas in Communications, Vol. 25, No. 3, pp. 589-600, Apr. 2007.

[6] H. Li, X. Feng, X. Gan, and Z. Cao, "Joint Specrtum Sensing and Transmission Strategy for Energy-Efficient Cognitive Radio Networks", CROWNCOM 2013, Washington DC, United States, July 2013.

[7] H. Liu, B. Krishnamachari, and Q. Zhao, "Cooperation and Learning in Multiuser Opportunistic Spectrum Access", ICC 2008, Beijing, China, May 2008.

[8] A. Raschellà, J. Pérez-Romero, O. Sallent, and A. Umbert, "On the use of POMDP for Spectrum Selection in Cognitive Radio Networks", CROWNCOM 2013, Washington DC, United States, July 2013.

[9] A. Raschellà, J. Pérez-Romero, O. Sallent, and A. Umbert, "On the impact of the Observation Strategy in a POMDP-based framework for Spectrum Selection", PIMRC 2013, London, UK, September 2013.

[10] K. Wang, L. Chen, Q.Liu, and K. Al Agha, "On Optimality of Myopic Sensing Policy with Imperfect Sensing in Multi-Channel Opportunistic Access", IEEE Transactions on Communications, Vol. 61, No. 9, pp. 3854-3862, September 2013.

[11] A. Raschellà, A. Umbert, J. Pérez-Romero, and O. Sallent, "A Testbed Platform to Demonstrate Spectrum Selection in Opportunistic Networks", WMNC 2013, Dubai, UAE, April 2013.

[12] L. Kleinrock, Queueing Systems. Volume I: Theory, John Wiley & Sons, 1975.

[13] http:// www.ettus. com.

[14] http://www.gnuradio.org.

[15] M. López-Benítez, and F. Casadevall, "Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio", European Transactions on Telecommunications (Wiley), Special Issue on European Wireless 2009 Conference, Vol. 1, no. 8, pp. 680-693, December 2010.